Behavioral Graphs

- Bayesian Knowledge Bases
  - Better than alternative
  - Verification & Validation
- Detecting Misinformation
  - Example
  - How it is a gray area of counterintelligence

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

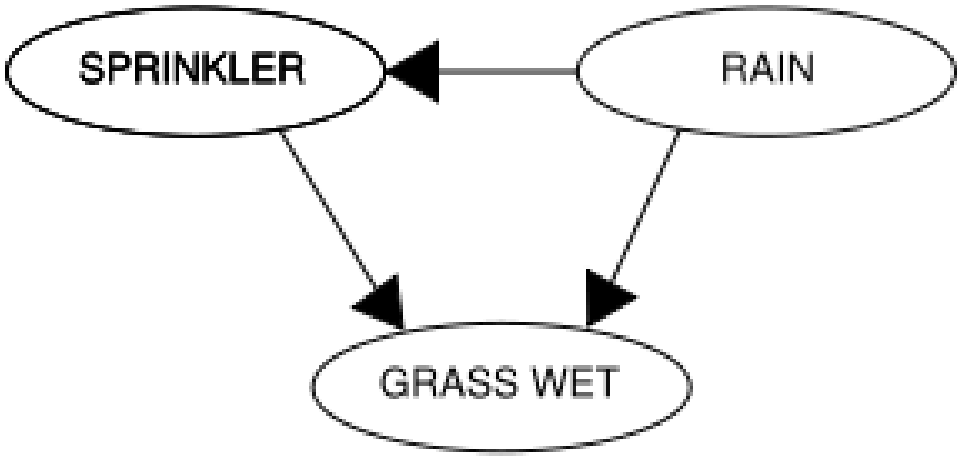
Graduate Student: Richard S. Detsch

Advisor: Prof. Eugene Santos Jr.

Bayesian Network

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99

2^2

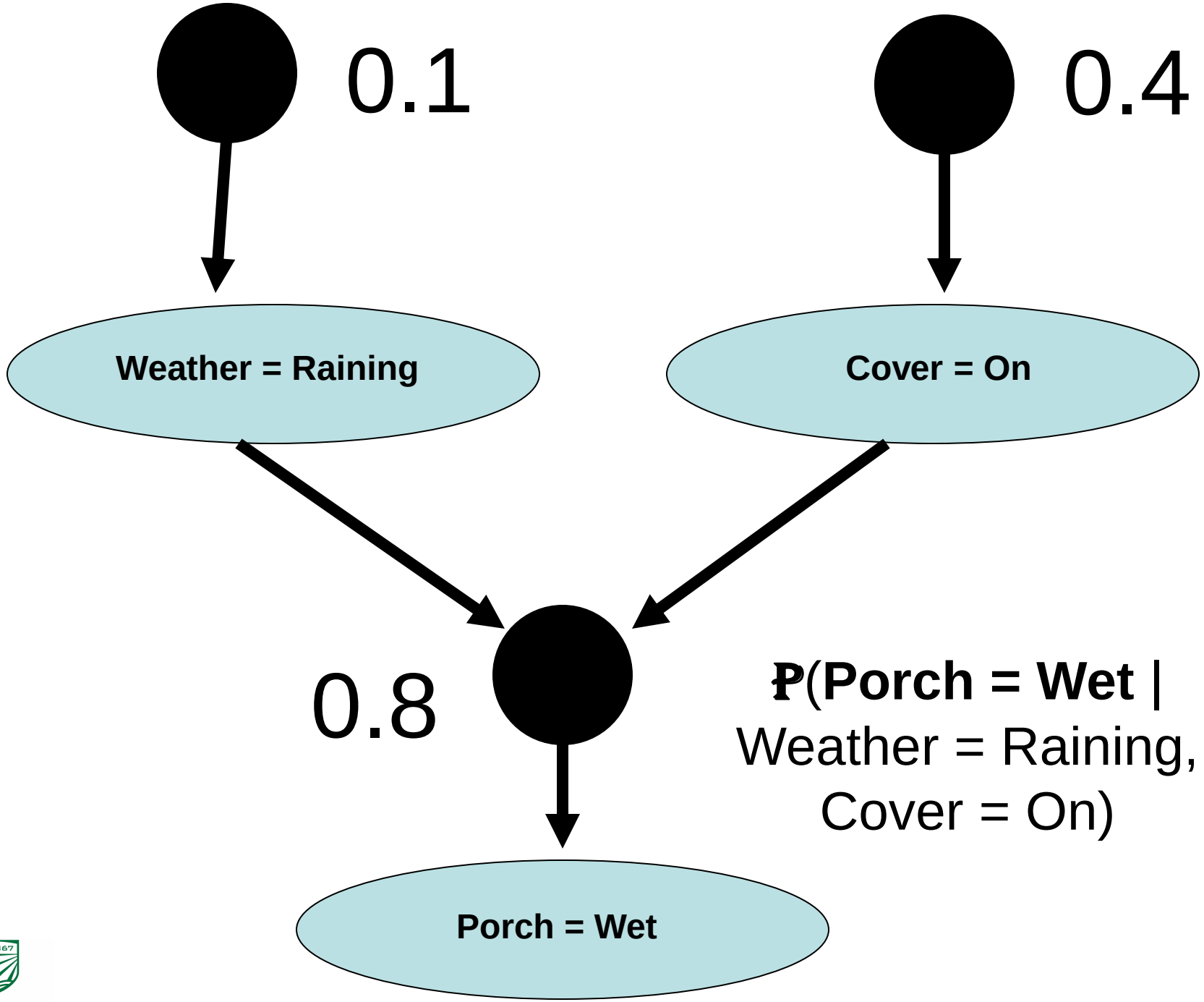


	RAIN	
T	F	
0.2	0.8	

2^1

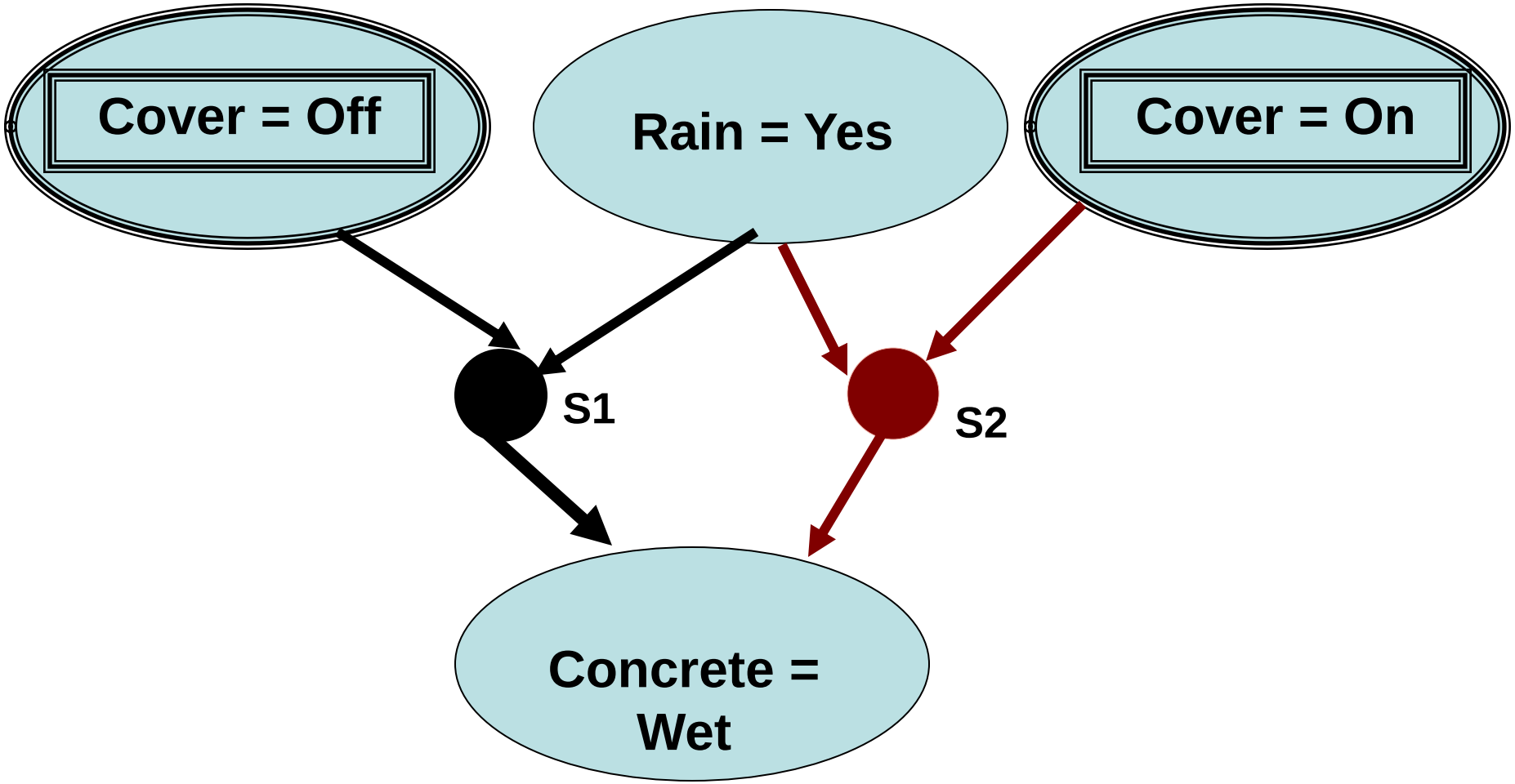
2^3

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01



P(**Porch = Wet** | Weather = Raining, Cover = On)

? Mutual Exclusion ?



North Korean **CAPABILITIES**

- Nuclear=Weak/Strong
- Army=Weak/Strong
- Air Force=Weak/Strong
- Navy=Weak/Strong

North Korean **INTENTIONS**

- Ask Seoul For Help
- Occupy Seoul
- Destroy Seoul

North Korean **BELIEFS** of South Korea beliefs of their intentions

- Ask Seoul For Help
- Occupy Seoul
- Destroy Seoul

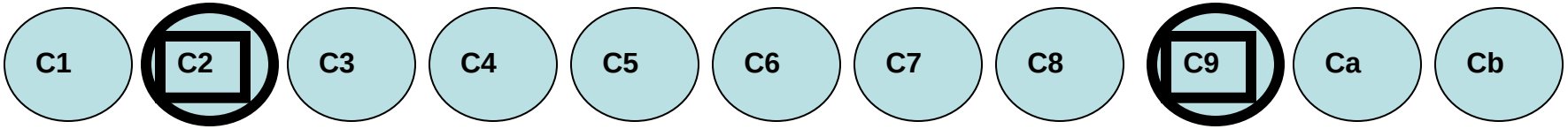
North Korean **OPPORTUNITIES**

- Regime Stable/Un-Stable
- Seoul will Attack/Not-Attack
- Military Sales Good/Bad
- Russia/China protection treaty passes
- Russia/China support increases/decreases

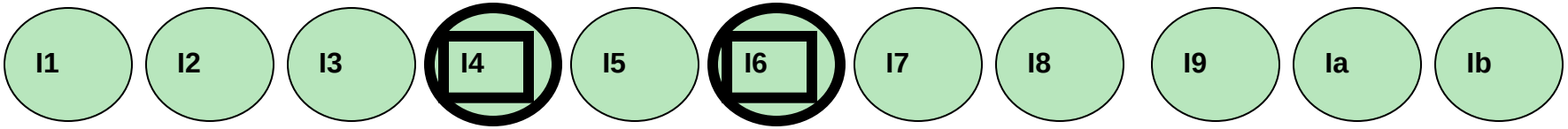
North Korean **ACTIONS**

- Ask Seoul For Help
- Occupy Seoul
- Destroy Seoul

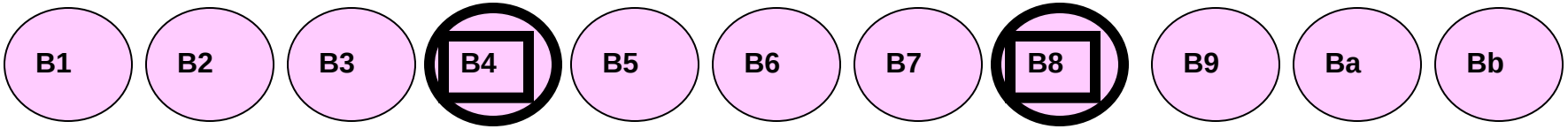
April 5th 2003: Capabilities :: Nuclear = Strong, Air Force = Strong



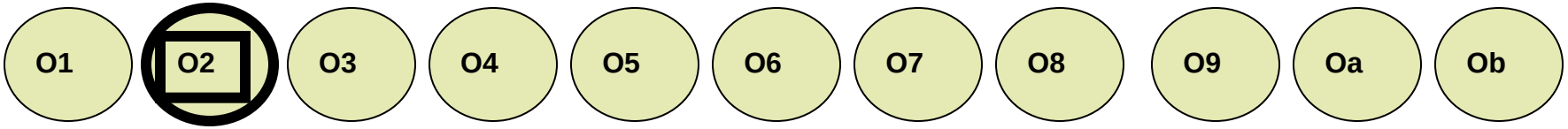
April 7th 2003: Intentions :: Completely Destroy Seoul, Occupy until 37th parallel



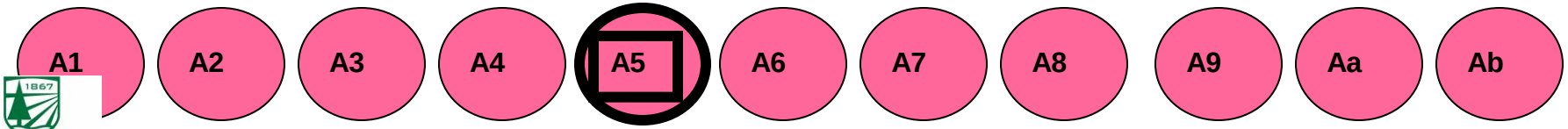
April 12th 2003: Beliefs :: **DPRK thinks** ➔ ROK thinks DPRK wants to Completely Destroy Seoul and use CBW on the DMZ

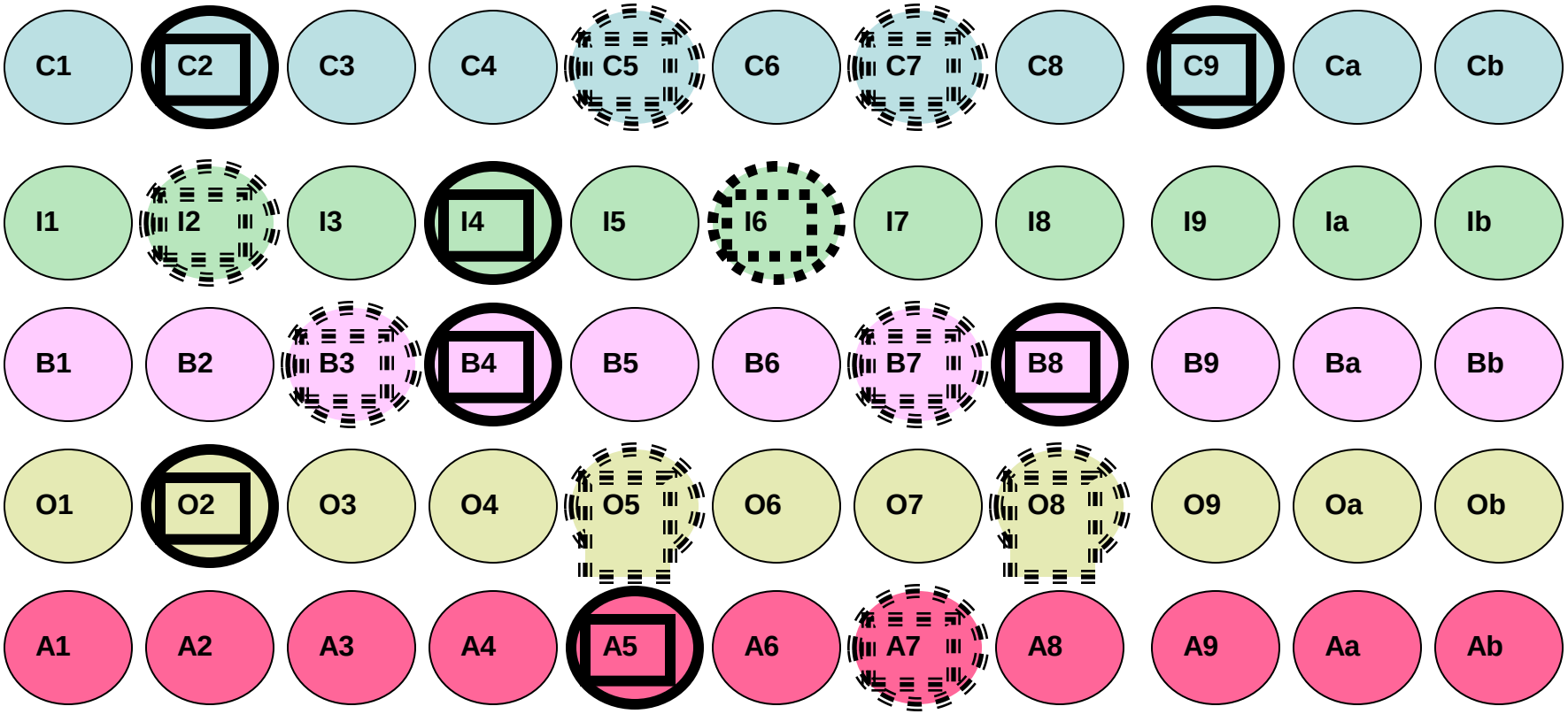


April 15th 2003: ROK decides to lift US imposed trade sanctions on DPRK



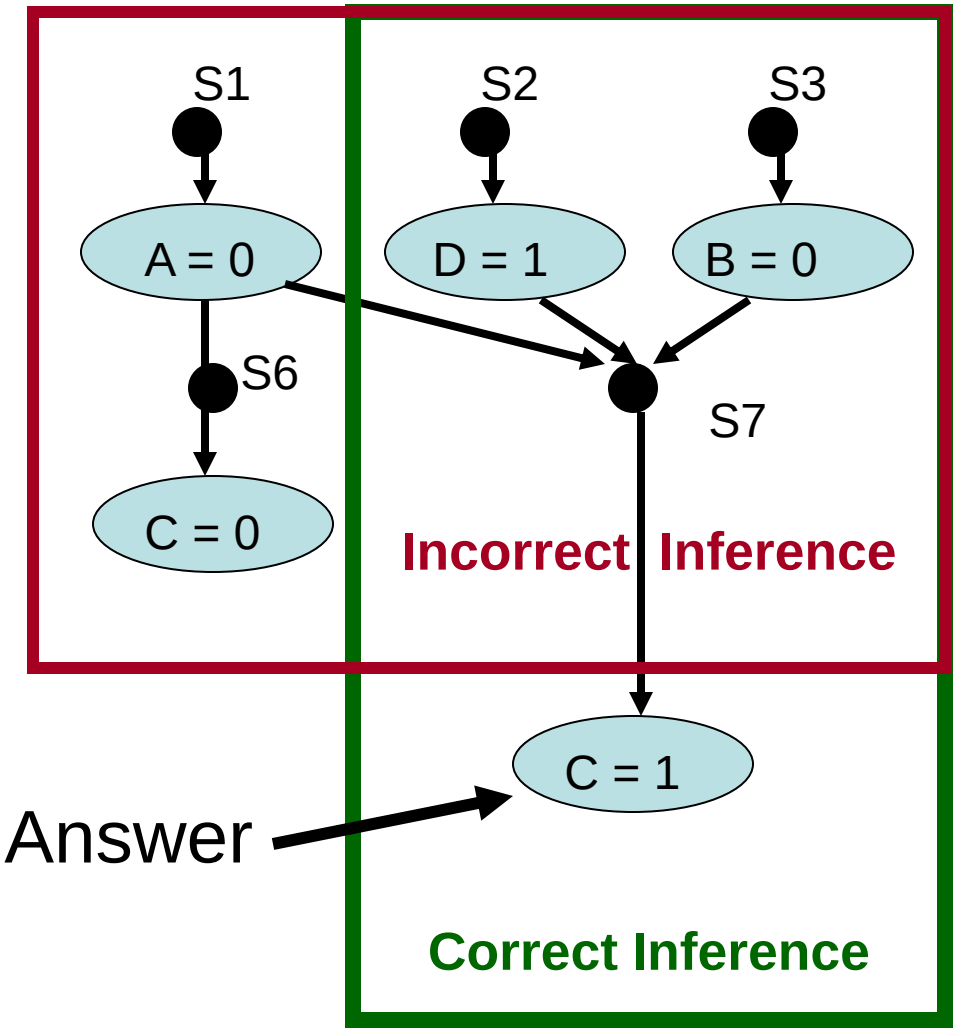
April 18th 2003: DPRK responds by initializing plans for the Kaesong Industrial Center

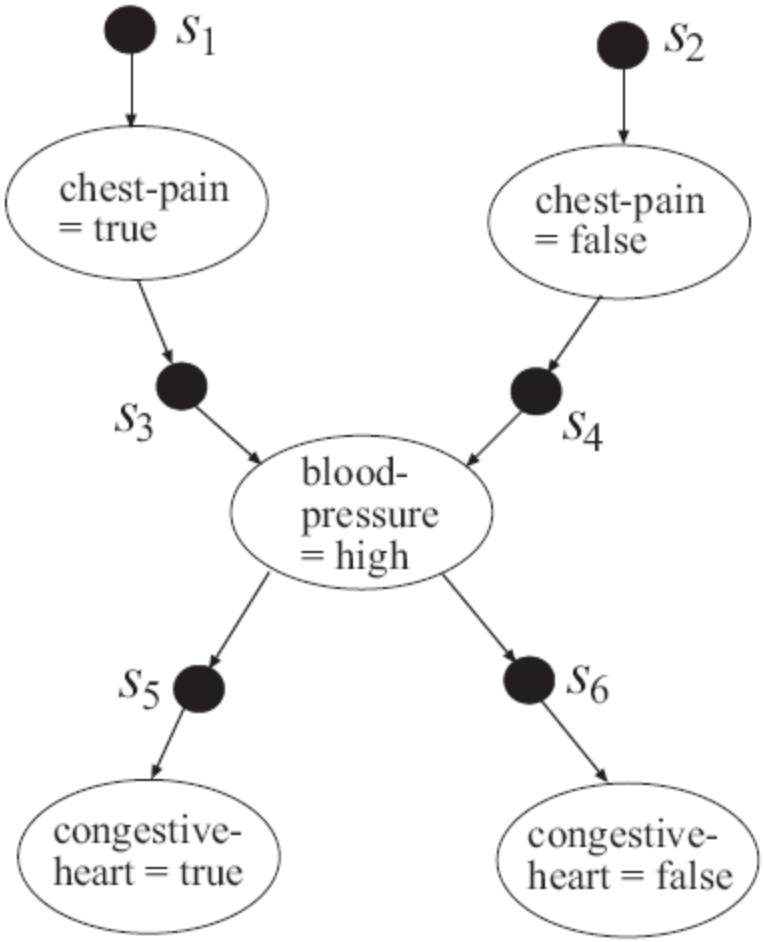


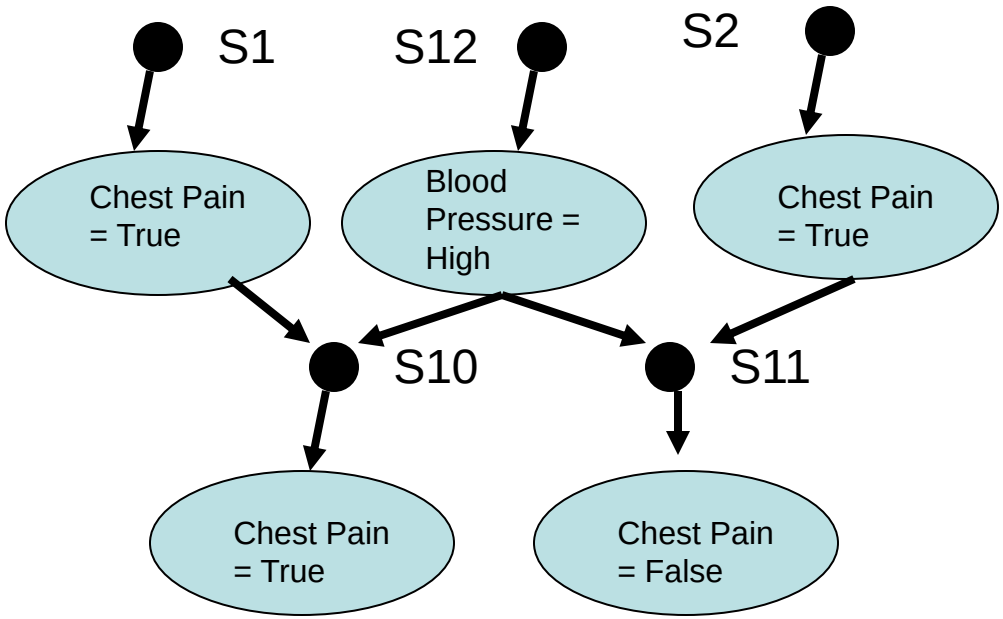


Correct Inference

- A **correct inference** for a test case is a complete state that contains the evidence, answer and has higher probability than any incorrect inference.
- An **incorrect inference** is a complete state that contains the evidence and a r.v. incompatible with the answer.





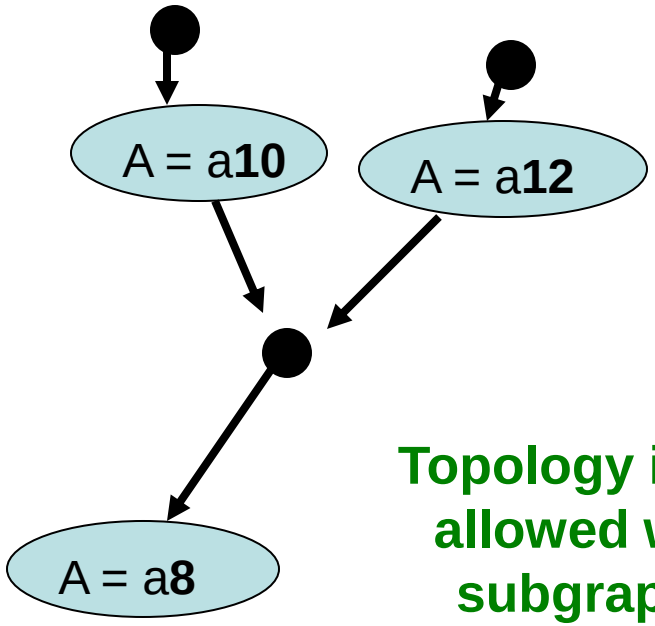


The graph on the right should replace the graph on the left, where S10 = S3*S5, S11 = S4*S6 and S12 = 1; and the way you can detect when thrashing takes place is whenever you have an I-Node with mutually exclusive antecedent S-nodes and consequent S-nodes whose consequents have more than on instantiation from a single random variable.

Topological Ordering

Let $\{c_1, \ldots, c_m\}$ be the I-nodes of an inference where $c_i \iff$ to $A_i = a_i$

Note: will not work for cycles in the graph!



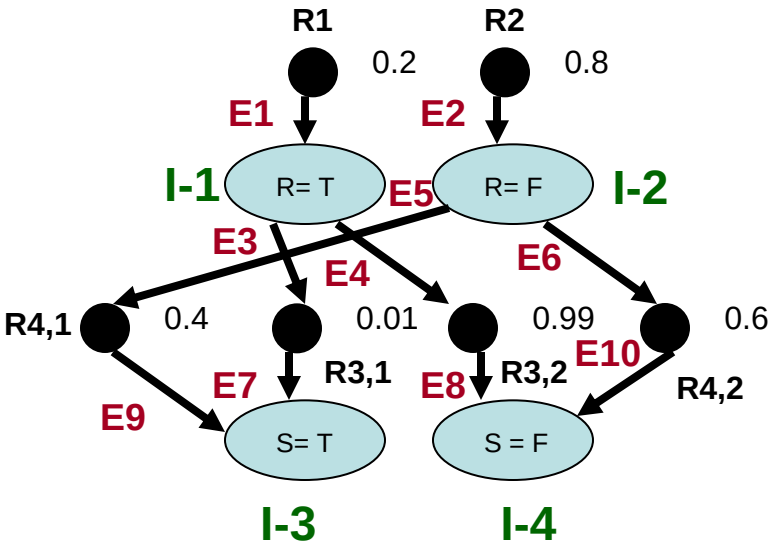
Good: 8 < min{10,12}

Topology is only allowed when subgraph is acyclic

*Then the **Probability** of the Inference would be:*

$$\prod_{i=1}^{m} P(A = a_i | A = a_{i+1}, \ldots, A = a_m)$$

Topological Ordering for Quasi-Unique Representation



Depth First?

What do you do about Cyclicity?

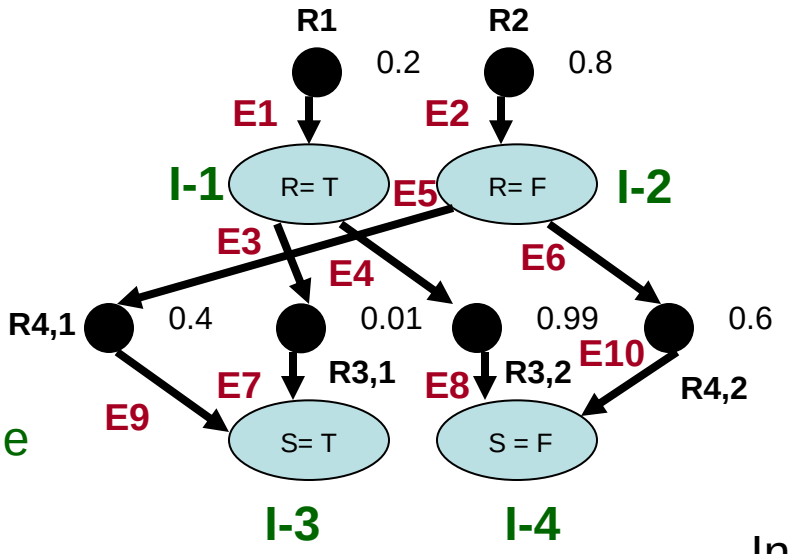
Have not formally decided …

On both vertices and arcs?

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Adjacency

1. Tailless S-nodes removed from columns.
2. Headless I-nodes removed from rows
3. S-node rows have exactly on element
4. I-node rows have one or multiple unity values



Incidence

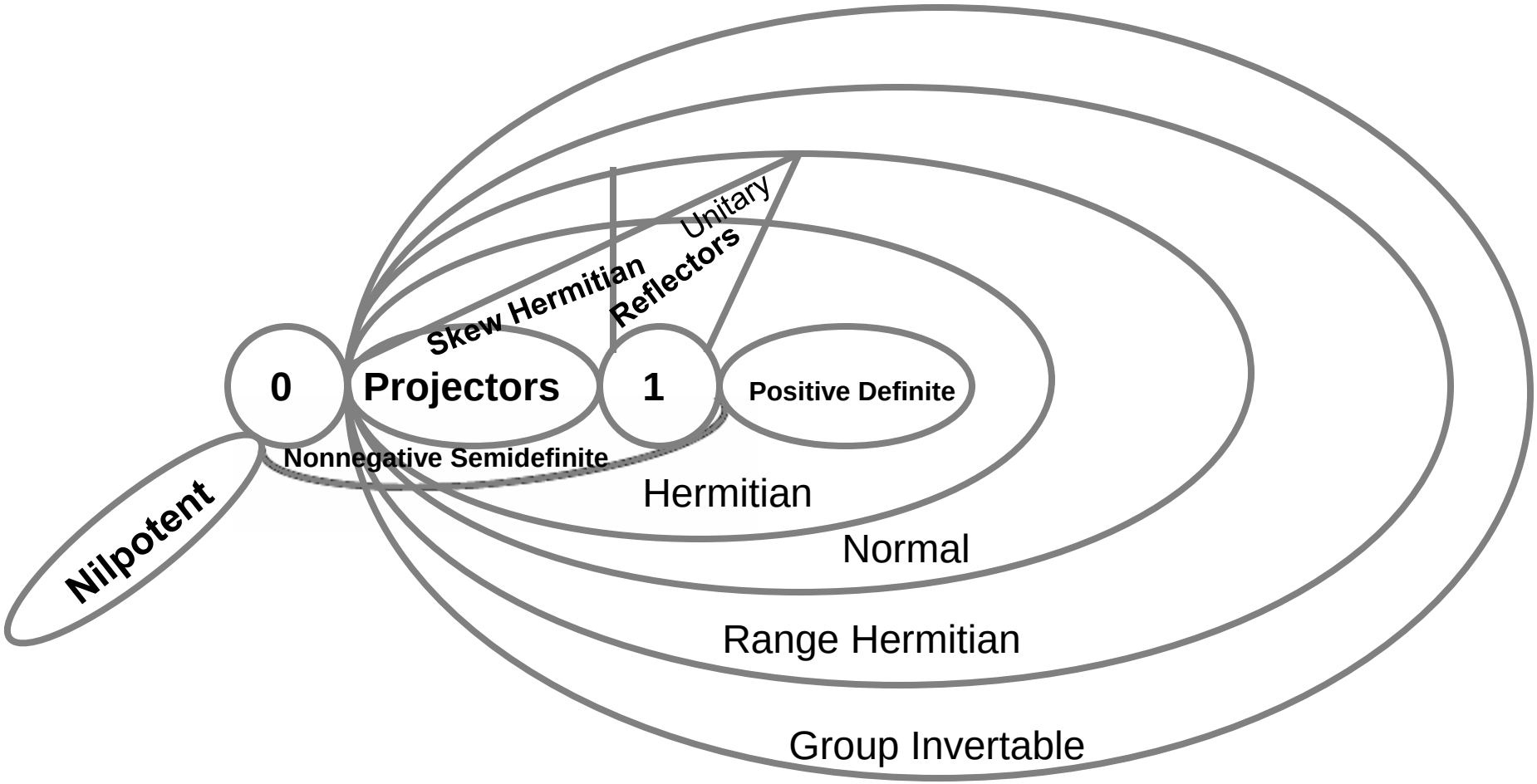
1. Two elements in every column: one positive one negative
2. For every row |negative| = # edges leaving, |positive| = # edges leaving

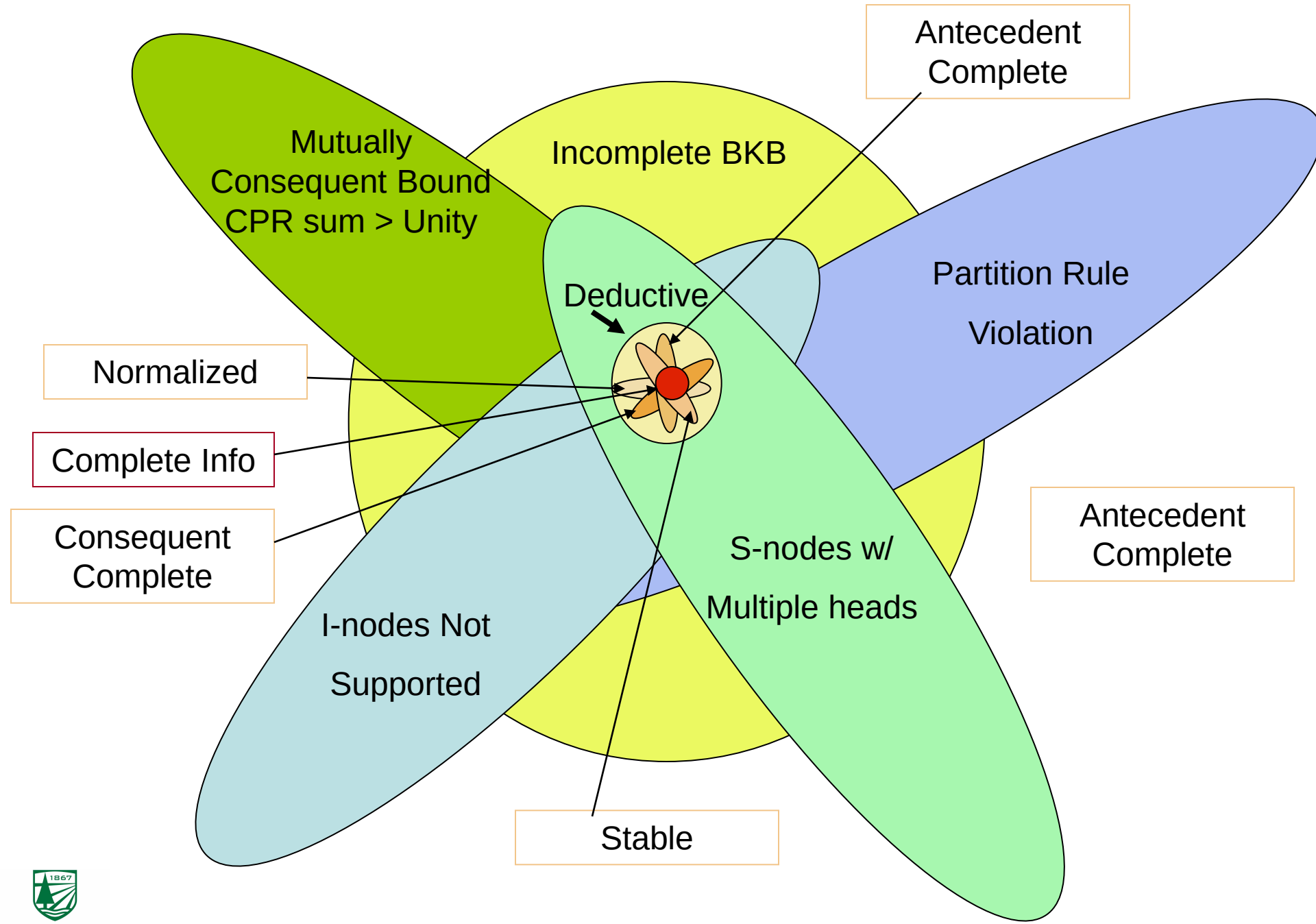
Adjacency Matrix

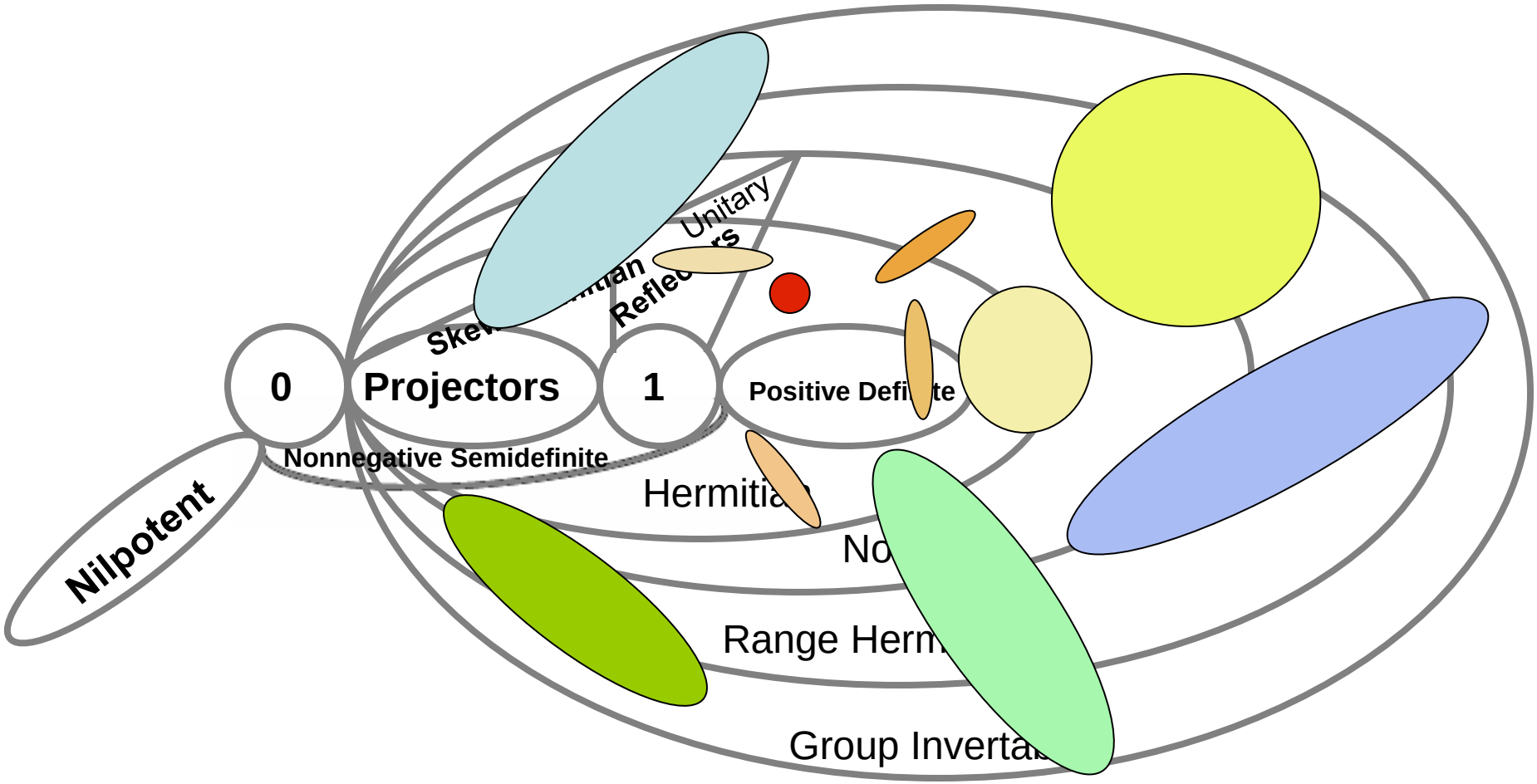
	I1	I2	R3,1	R3,2	R4,1	R4,2	I3	I4
R1	0.2							
I1			1	1				
R2		0.8						
I2					1	1		
R3,1							0.01	
R3,2								0.99
R4,1							0.4	
R4,2								0.6

Incidence Matrix

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
R1	-0.2	0	0	0	0	0	0	0	0	0
I1	0.2	0	-1	-1	0	0	0	0	0	0
R2	0	-0.8	0	0	0	0	0	0	0	0
I2	0	0.8	0	0	-1	-1	0	0	0	0
R3,1	0	0	1	0	0	0	-0.01	0	0	0
R3,2	0	0	0	1	0	0	0	-0.99	0	0
R4,1	0	0	0	0	1	0	0	0	-0.4	0
R4,2	0	0	0	0	0	1	0	0	0	-0.6
I3	0	0	0	0	0	0	0.01	0	0.4	0
14	0	0	0	0	0	0	0	0.99	0	0.6







Misinformation

Conclusion/Q.A.

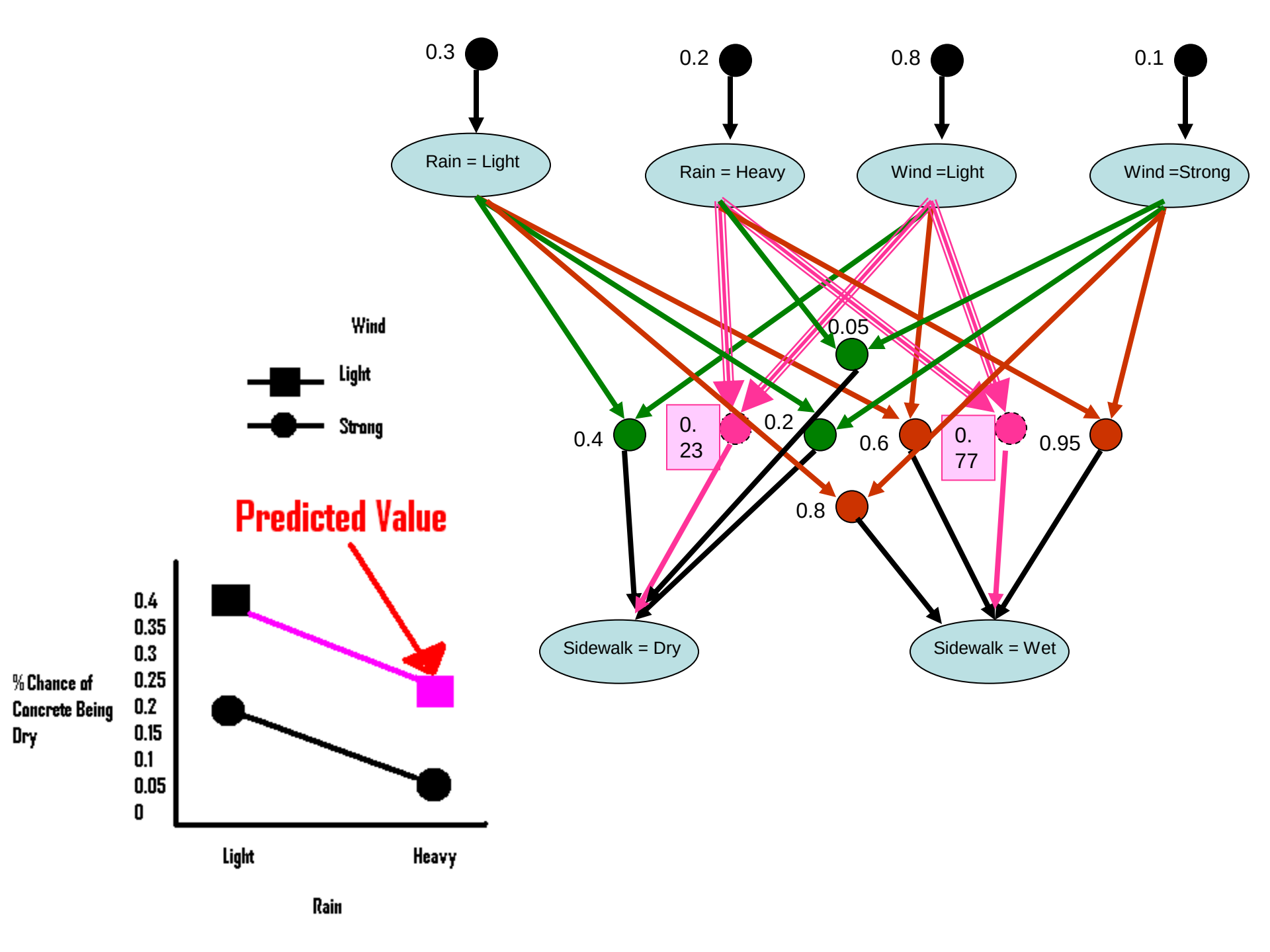
- How to create, store, structure and query Bayesian Knowledge Bases.
- How to detect when a US intelligence analyst is engaging in misinformation given the papers/products she writes, how she queries, her electronic dialog and even biometrics.

Annex

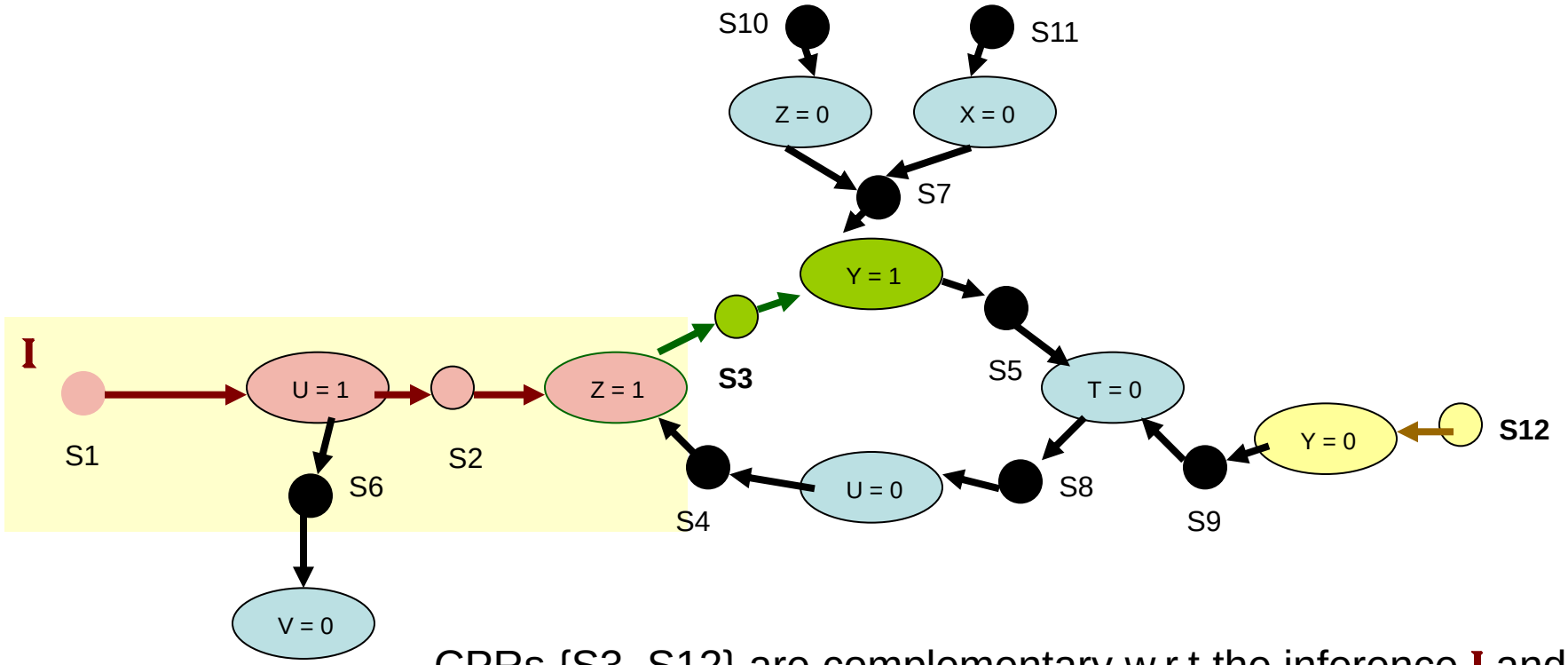
- Filling in missing data for a Bayesian Knowledge Base.
- Cyclical Knowledge

Sidewalk Wetness Knowledge

- Test Cases
  - P(Sidwalk = Dry | Wind = Light, Rain = Light) = 0.4
  - P(Sidwalk = Dry | Wind = Strong, Rain = Light) = 0.05

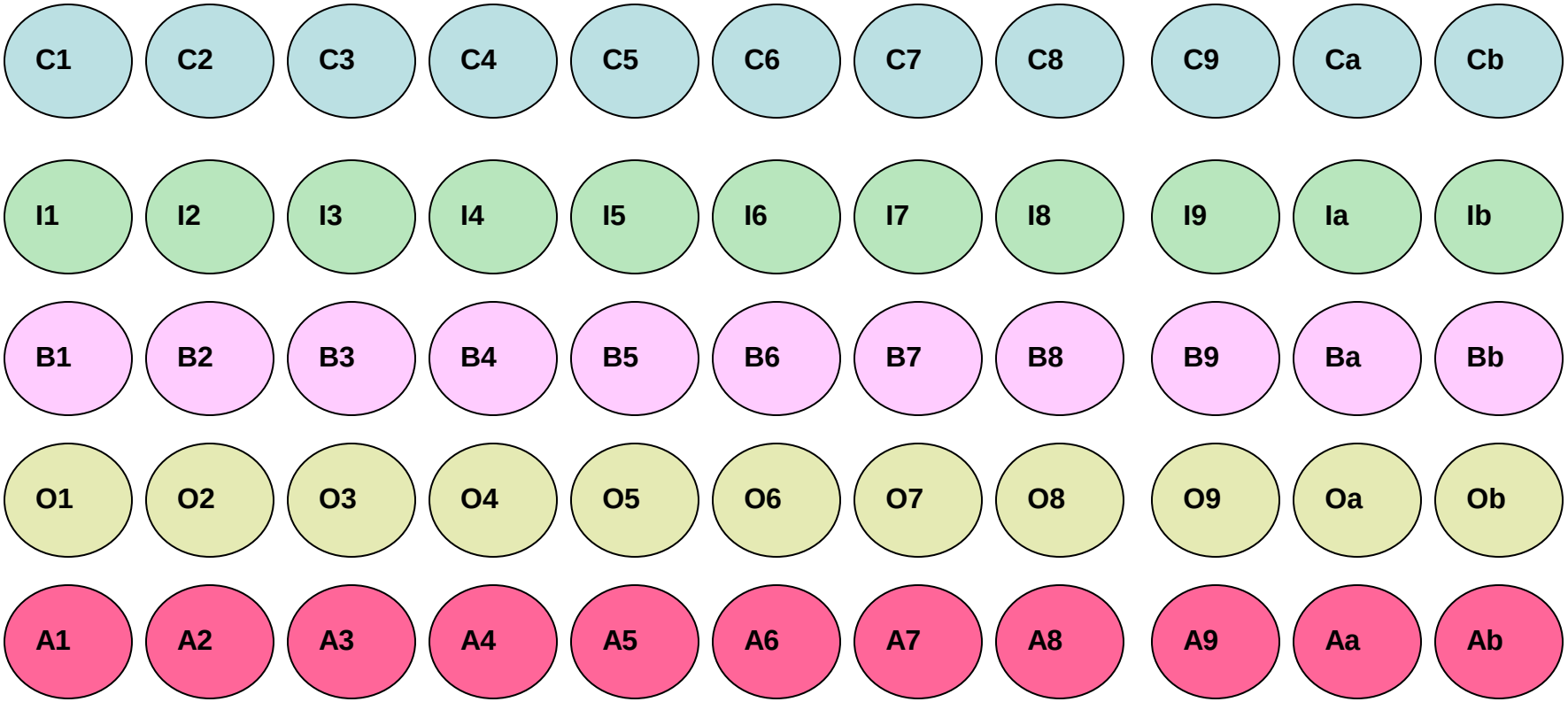

0.3
0.2
0.8
0.1
Rain = Light
Rain = Heavy
Wind =Light
Wind =Strong
0.05
0.4
0.
23
0.2
0.6
0.
77
0.95
0.8
Sidewalk = Dry
Sidewalk = Wet
Wind
Light
Strong
Predicted Value
% Chance of Concrete Being Dry
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Light
Heavy
Rain

A set of CPRs is **complementary** w.r.t. an inference and a r.v. if each extents the inference by including a unique instantiation of the r.v.



CPRs {S3, S12} are complementary w.r.t the inference **I** and the r.v. Y additionally if Y=0 and Y=1 were the only instantiations for Y then {S3, S12} is the unique maximal complementary set of CPRs.

Extras


C1 C2 C3 C4 C5 C6 C7 C8 C9 Ca Cb
I1 I2 I3 I4 I5 I6 I7 I8 I9 Ia Ib
B1 B2 B3 B4 B5 B6 B7 B8 B9 Ba Bb
O1 O2 O3 O4 O5 O6 O7 O8 O9 Oa Ob
A1 A2 A3 A4 A5 A6 A7 A8 A9 Aa Ab

Concerns

- Is it necessary to have 100% assurance of all test cases or is having a temporal priority enough, for example if you have two test cases at separate times with equivalent evidence and contradictory answers, this could just simply mean the groups behavior has changed